U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 277

# Some Performance Comparisons Between NMC's Spectral Model and the Seven-Layer Primitive Equation Model

Donald G. Marks
John D. Stackpole
Development Division

May 1983

Some Performance Comparisons Between
NMC's Spectral Model and the
Seven-Layer Primitive Equation Model

by

Donald G. Marks
John D. Stackpole

ABSTRACT

The performance of NMC's new global spectral model is compared to the performance of the older, seven-layer, grid-point primitive equation model which it replaced. Since it has been in operation since August 1980, mean error statistics may be used to determine this performance. In general, the spectral model shows a distinct improvement in overall forecast skill. The spectral model especially shows an improvement in wave amplitudes and generally stronger flow. Much of this improvement occurs in the longer waves and leads to greatly improved medium range forecasts.

## 1. INTRODUCTION

On August 12, 1980, the National Meteorological Center (NMC) initiated operatonal use of a spectral model. This ended the long and glorious career of the primitive equation (PE) model first detailed by Shuman and Hovermale in 1968. The new spectral model is explained by Sela (1980) and will not be discussed at length here. Among the salient features included in both models are the following: large-scale and convective moisture processes, evaporation, and sensible heating over oceans, orography and surface friction. There is presently no radiation parameterization in the spectral model. The initial fields come from the operational (Hough) analysis for both, but then are altered by a normal mode initialization prior to the actual forecast for the spectral model. The spectral model is also a global model whereas the PE model has hemispheric (although the spectral model was hemispheric past the 48-hour forecast until March 18, 1981.)

As with any major change in the NMC operational products, questions arise as to how the models differ in their forecasts. Of course, not all aspects of the forecasts can be evaluated in a single paper. In fact, all aspects have not even been investigated. We will therefore limit ourselves to presenting some of the more interesting areas of comparison.

## 2. Forecasts versus Analyses

Numerical weather forecasts have been objectively verified routinely at NMC since 1962. Various model forecasts are verified against the corresponding analysis. A continuous record of these verifications then exists which may be used to establish differences between the models.

The longest existing record is for the conventional skill (S1) score (Teweles and Wobus, 1954). This S1 score is calculated on a 49 point grid-a subset of a 10° longitude by 5° latitude grid-covering the U.S. The forecasts are verified using the operational (Hough) analysis, the details of which are found in Flattery (1972). It is important to realize that the verifying area is over a data rich region where the analysis can be expected to be very accurate. Since the spectral model is also used to create the "First Guess" for the Hough analysis, S1 scores for short forecasts over ocean areas are suspect and will not be presented here.

The mean sea level and 500 mb S1 scores calculated from the Spectral model and the 7L PE are shown in figures 1 and 2. Since these are for different years, direct comparison between the two models is not possible. However, the Spectral Model does exhibit a noticable improvement in the early hours of the forecast. This is probably due to the utilization of the normal mode initialization technique used in that model. The Spectral Model also shows a wider variation of scores, with extremely good S1 scores in the winter, and rather poor scores in the summer. While this pattern was also true for the 7L PE, the spectral model seems to exagerate the tendency. As figure 3 shows, the winter months tend to dominate and the Spectral Model shows an overall yearly improvement in skill. Improvement in the summer scores will probably be dependent on further research in convection, radiation and other physical processes currently being done at NMC.

## 3. VERIFICATION AGAINST OBSERVATIONS

A direct method of forecast verification has also been in use at NMC for some time; we verify the forecasts against reported (radiosonde) observations and complete summary statistics, averaging over either (or both) time and space.

The procedure is quite straightforward - the forecast information (heights, temperature, relative humidity or winds) is interpolated (bi-linearly) from the model gridpoints to the location of the upper air observations, the difference (forecast minus observed) taken as the error of the forecast, and those errors, and squared errors are summed for as many forecast verifications as one wishes. We end up with the makings of the bias and RMS error pigeonholed at each observation location, at selected numbers of pressure levels, for the four forecast quantities.

Prior to the error calculation the observed data are compared to an analysis; any observations that are "sufficiently" different from the analysis are themselves considered to be in error and are excluded from the forecast error determination and the associated summations.

Once a sufficient number of forecast/observation verifications have been accumulated (say for a month) they can easily be presented as monthly mean verifications for individual stations or agglomerated into networks to give the space mean verifications for various geographical areas.

The system can also be used to verify single forecasts either at individual stations or over areas; this is usually done in the context of comparing two forecast models for their relative merits.

A number of station network areas have been assembled for various purposes - of primary interest here are two: "NH102"-102 stations quasi-uniformly distributed over the northern hemisphere (these stations are also used by the Air Force and Navy for their verifications, thus allowing (partial) intercomparison between models) - and "NA110": virtually the entire North American radiosonde network between 25° and 60° North.

Our longest continuous record is in Figure 4 showing the monthly root mean square absolute vector error of the 48 hr forecast wind at 500 mb for the northern hemisphere network. It starts with the 6L PE (1 bedient mesh) changes to the 7L PE (1/2 bedient mesh) in January of 1978, and to SMG3C (30 wave resolution) in August of 1980. Also shown are the Air Force and Navy model forecasts in parallel verification. (Those models have not changed appreciably over the years on the chart).

The figure speaks for itself, showing the substantial improvement of the 7L over the 6L model and no particular improvement with the introduction of the spectral model. (This latter was not expected based on earlier preimplementation tests).

Fig. 5 on the other hand does show an improvement in the S1 (gradient error) score at 250 mb since the spectral model began. This S1 score is computed between observation stations rather than specific grid points in an analysis; thus the numerical values are not strictly comparable in these figures and S1 scores elsewhere in this paper.

Fig. 6 shows a rather dramatic (and anticipated) improvement in the 250 mb speed bias for the winter of 80-81 vis a vis 79-80 (and to a lesser extent 78-79). This is attributable to the increased vertical resolution of the spectral model and particular attention being paid to internal viscosity calculations.

Fig. 7 demonstrates a (slight) improvement in the 24 hour temperature forecasts at 250 mb since the inception of the spectral as measured by the rms error. The analysis persistence error is also shown for reference and is quite stable within the limits of the seasonal variations.

The next two figures, #8 and #9, take us to 100 mb and both show potential, if minor, problems. The temperature bias shows the spectral to be running colder than the 7L PE was at these higher levels. At present the spectral lacks any radiation computations (the 7L PE had rudimentary ones); presumably when such are introduced the bias will tend toward zero.

The speed bias - spectral consistently too fast where the 7L PE was generally too slow - most likely arises from the relative lack of resolution in the spectral model stratosphere. It is not likely that this can be overcome until we obtain the next generation computer.

Other statistics at other pressure levels and for other station networks have been obtained - the results from them are consistent with what has been shown here as far as the relative strengths of the spectral model vs the 7L PE are concerned.

4. Wave Number Behavior

NMC has been verifying the operational model by wave number for only a relatively short time. This, of course, greatly limits the conclusions possible when comparing the two models. However, certain statistics are available by wave number from 1979 until the present. Examining these wave number statistics indicates differences between the models which are not obvious from the grosser statistics of S1 scores. Previous authors who have compared models by wave number include Pratt (1979), Houghton and Irvine (1976) and Baumhefner and Downey (1978).

The analysis used for comparison with the two models is the "Hough" described by Flattery (1971). Although the Spectral model and the Hough analysis are global, wave oriented products, the 7L PE was a grid point model. Therefore, in order to have comparable statistics, all values are calculated off the polar stereographic map. This is the normal, operational product, 381 km between grid points at 60°N. These grid point values are then interpolated to points spaced 2.5° latitude by 5° longitude. At each constant latitude a harmonic analysis is then performed, yielding zonal wave amplitudes and phase angles.

While any field may be studied this way, only the height field (on constant pressure) will be referred to in this paper. We first consider amplitude error, calculated by averaging the daily amplitudes of the observed and forecast waves. The averages of these amplitude errors are shown in figures 10 and 11 for the years 1979 and 1981 respectively. These figures only show one latitude (45°N), level (500 mb) and forecast time (36-hr), but are represrentative of the performance of the two models.

Until now, it has been almost axiomatic that numerical forecast models underforecast the amplitude of waves. The 7L PE definitely exhibits this tendency, but the Spectral model shows a vast improvement in forecasting wave amplitudes correctly. This improvement in forecasting amplitudes does not show up in the S1 scores, since these scores are more sensitive to phase angle errors than to amplitude errors (phase angle errors do not seem to differ significantly). This change in the model's behavior could have far-reaching implications to forecasters who use the model output to forecast type and quantity of significant weather.

The variability of the height field can also be investigated by wave number. At present, this is done by calculating the RMS error (by wave number groups) over the northern hemisphere (from 20°N to 90°N) and from 1000 mb to 100 mb. When compared to the natural variance of the atmosphere, this gives a single statistic that may be used to infer total model performance by wave group. Without going into elaborate details, we can say that there is virtually no difference between the two models for wave number groups 4-9 or 10-15. For wave groups 1-3, however, the spectral model shows a definite improvement in forecasting skill. In fact, during 1981 the Spectral model had an rms error for waves 1-3 amounting to 33% of the average standard deviation of the analysis fields. The 7L PE model, however, had an error of 38% during its last year as the operational model. This improvement in forecasting the very long waves may help explain some of the longer range forecast results explained in the next section.

## 5. Medium Range Forecasts

NMC has been making 3-5 day forecasts for several years. Until recently, however, these tended to be man-made (as opposed to computer model) products. Therefore, objective comparisons between the performance of the 7L PE and the Spectral model are very limited. In addition to the subjective evaluation by the forecasters, the 500 mb standardized correlation scores for this 3-5 day period seem to reflect the improvement of the Spectral over the 7L PE. This score improved from 48 in 1979 to 54 in 1980 and to 56 in 1981. In fact, the Spectral model shows some skill out to 7-8 days. The 7L PE model, since it took longer to run, was not usually run out this far, so comparisons are not possible. However, the fact that the Spectral model, which is a global model, shows skill in these ranges does tend to substantiate Somerville's (1980) conclusions regarding the importance of the tropics and cross equatorial flow. At NMC, the number of waves used and the number of levels used in the Spectral model at longer time periods are both reduced, but the model remains global during the entire forecast.

6. Conclusions

The Spectral model is a definite improvement over the 7L PE model, although it may not be immediately obvious. Short range S1 scores do not change much. However, the spectral model definitely maintains more energy as shown by the wind speeds and wave amplitudes. This higher energy and better long wave forecast enables the model to show a definite improvement at longer forecast times (3-5 days). The main differences in the short range forecasts (poor S1 scores in summer and cooler temperatures at upper levels) seem to be related to physical processes and radiation. These should improve as new techniques are developed to complement the excellent handling of the dynamics in the model.

## REFERENCES


Baumhefner, D. and P. Downey, 1978: Forecast intercomparisons from three numerical weather prediction models. MWR, 106, pp. 1245-1279.

Flattery, T. W., 1971: Spectral models for global analysis and forecasting. Air Weather Service Tech. Rpt. 242, pp. 42-54.

Houghton, D. D. and W. S. Irvine, 1976: A case study comparison of the performance of operational prediction models used in the U.S. MWR, 104, No. 7, pp. 817-827.

Pratt, R. W., 1979: A space-time comparison of the NCAR and GFDL general circulation models of the atmosphere. J. Atmos. Sci., 36, pp. 1681-1691.

Sela, J. G., 1980: Spectral modeling at the National Meteorological Center. MWR, 108, pp. 1279-1292.

Shuman, F. and J. Hovermale, 1968: An operational six-layer primitive equation model. J. Appl. Meteor., 7, 525-547.

Somerville, R. C. J., 1980: Tropical influences on the predictability of ultralong waves. JAS, 37, No. 6, pp. 1141-1156.

Teweles, S. and H. Wobus, 1954: Verification of prognostic charts. Bull. Amer. Meteor. Soc., 35, pp. 455-463.

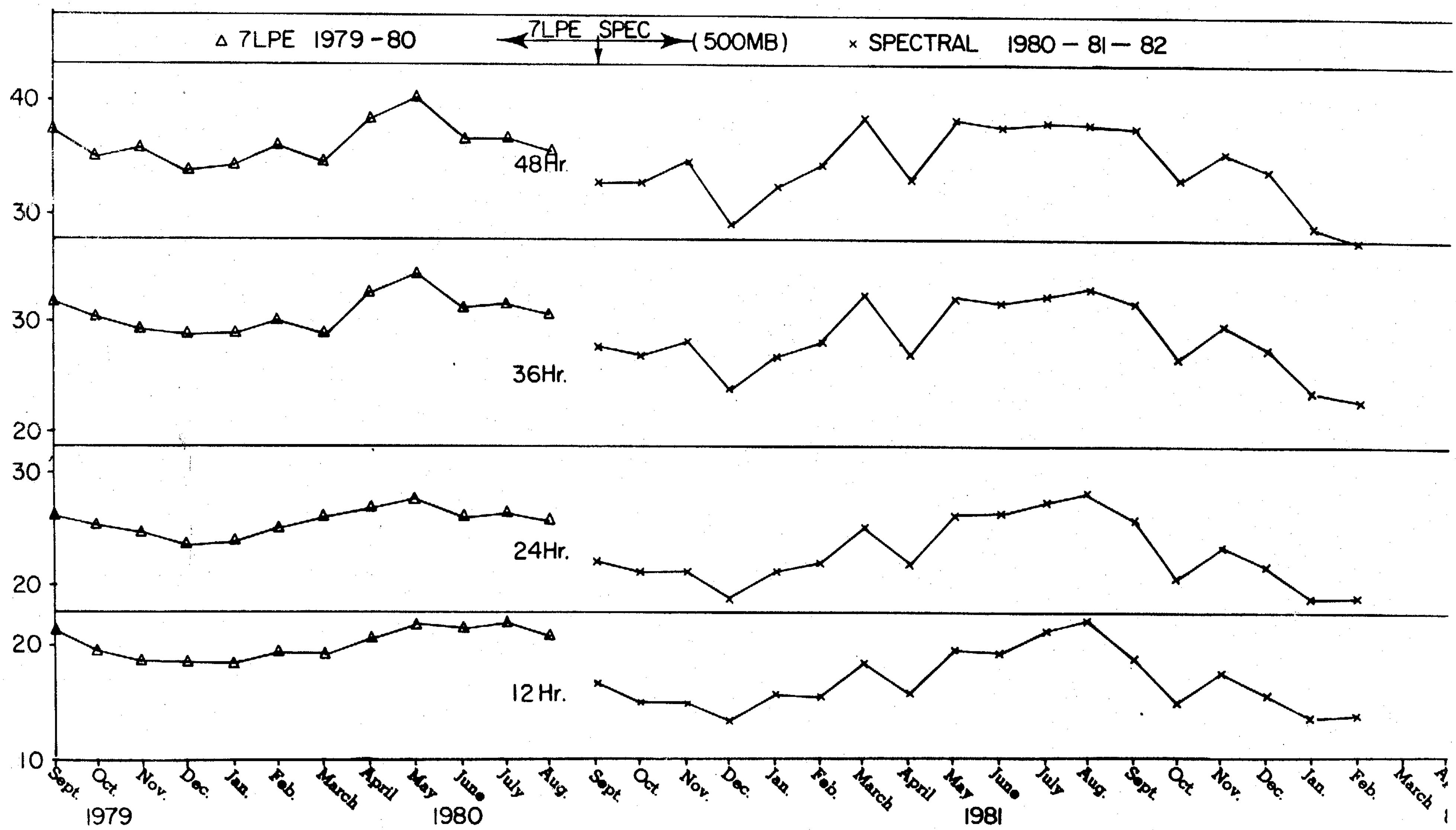



FIGURE 1

MONTHLY MEAN S1 SCORES FOR 7L PE AND SPECTRAL MODEL

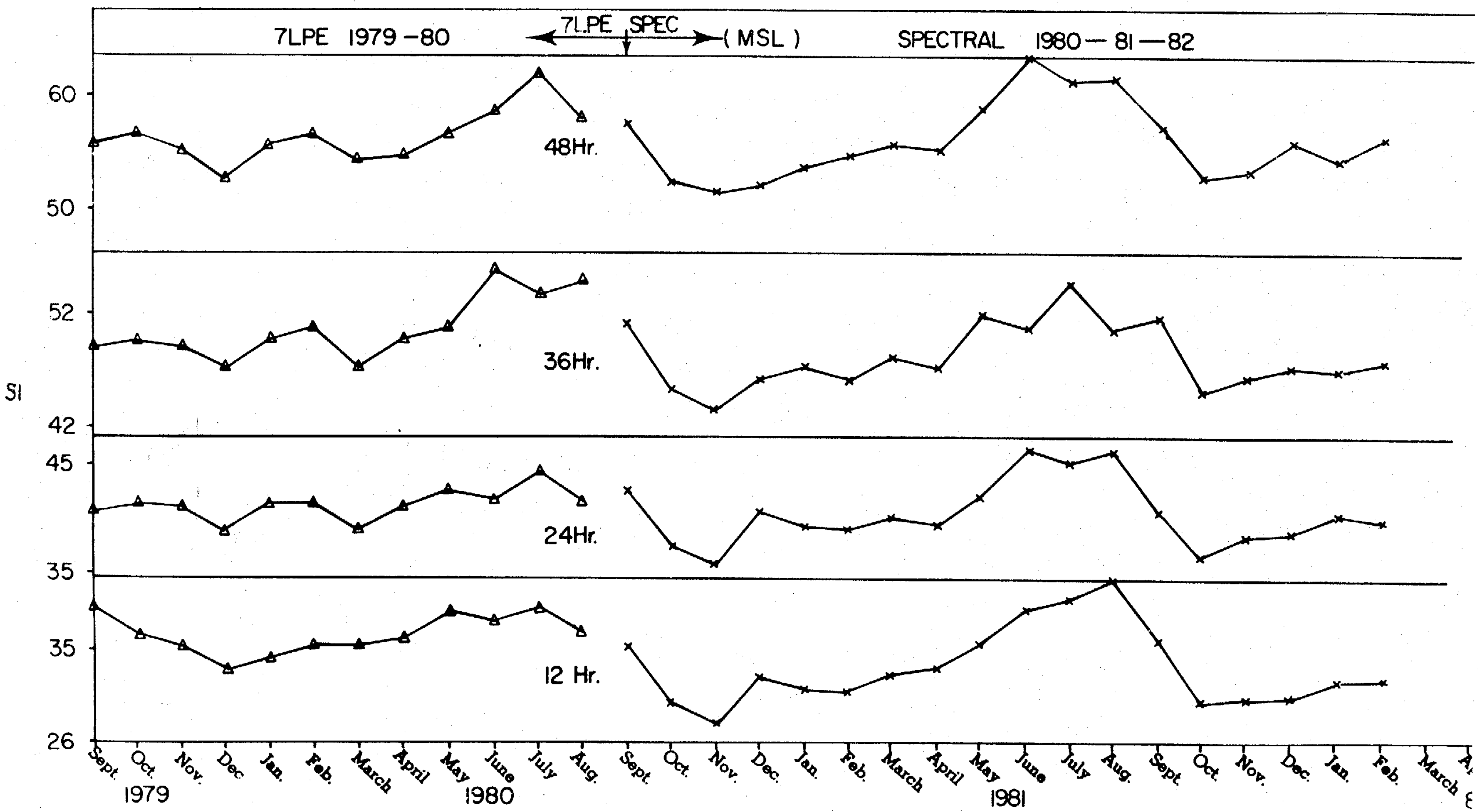
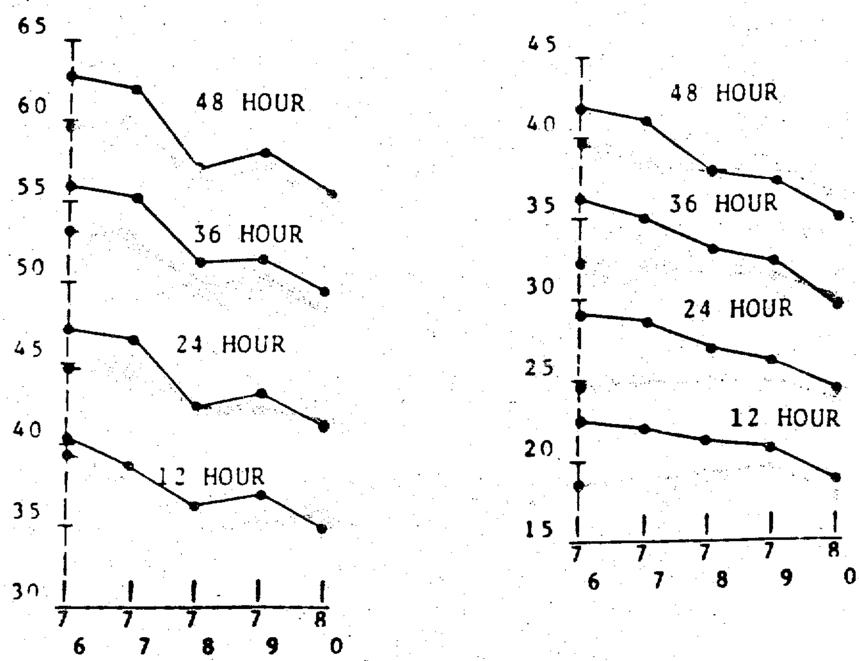



FIGURE 2

MONTHLY MEAN S1 SCORES FOR 7L PE AND SPECTRAL MODEL

S1 COMPARISONS OF THE SPECTRAL AND LFM FOR THE CONTIGUOUS UNITED STATES. GRID USED: 49-POINT LAT/LONG GRID. THIS IS A SUBSET OF A 63-POINT GRID WHICH COVERS THE AREA BETWEEN 65 WEST AND 145 WEST LONGITUDE, AND BETWEEN 25 NORTH AND 55 NORTH LATITUDE. GRIDPOINT SPACING IS 5 DEGREES LATITUDE BY 10 DEGREES LONGITUDE.

——— SPECTRAL




Sea Level

500 mb

Figure 3

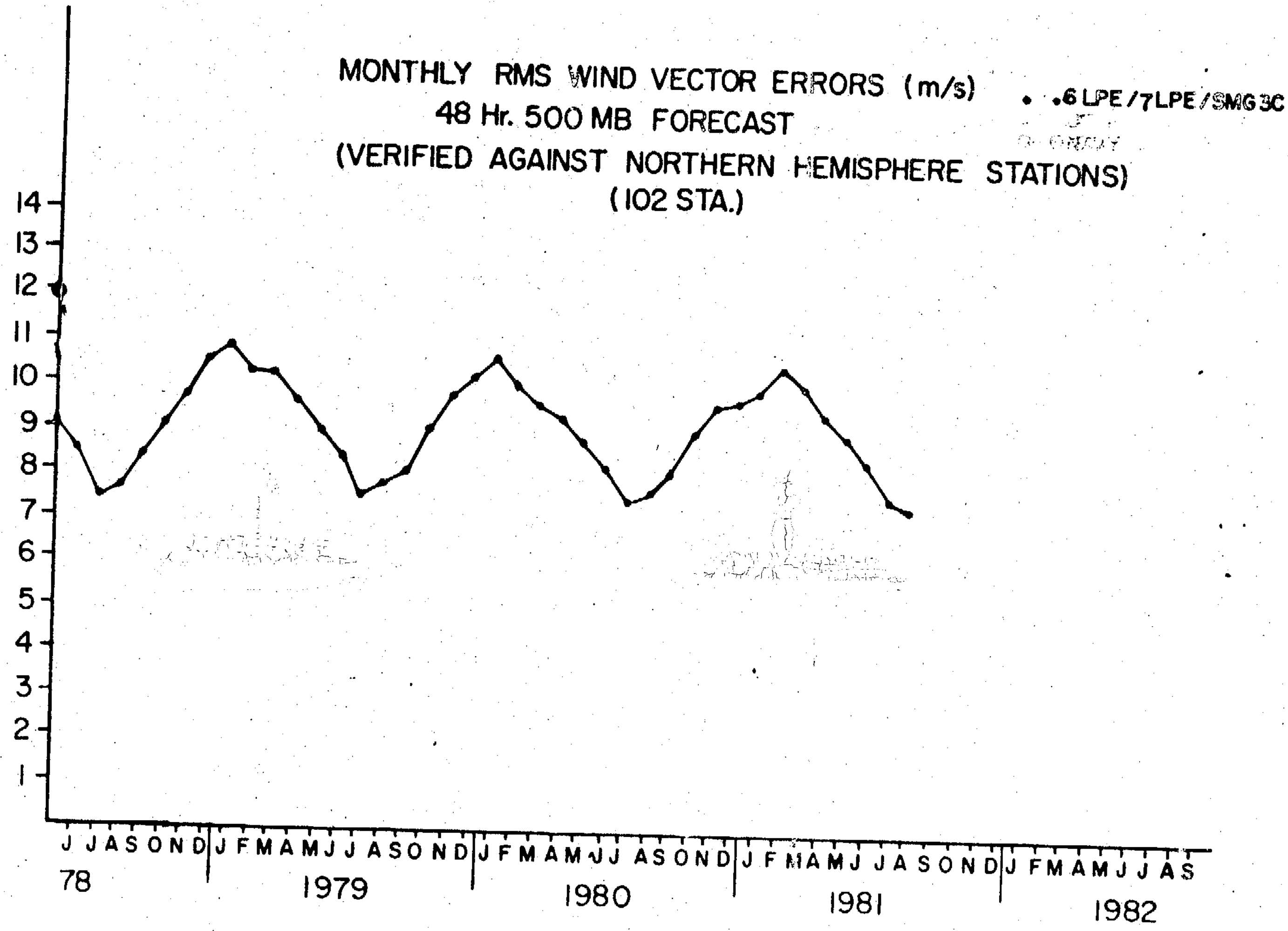

MONTHLY RMS WIND VECTOR ERRORS (m/s)
48 Hr. 500 MB FORECAST
(VERIFIED AGAINST NORTHERN HEMISPHERE STATIONS)
(102 STA.)
6LPE/7LPE/SMG3C
14
13
12
11
10
9
8
7
6
5
4
3
2
1
J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S
78
1979
1980
1981
1982

FIGURE 4

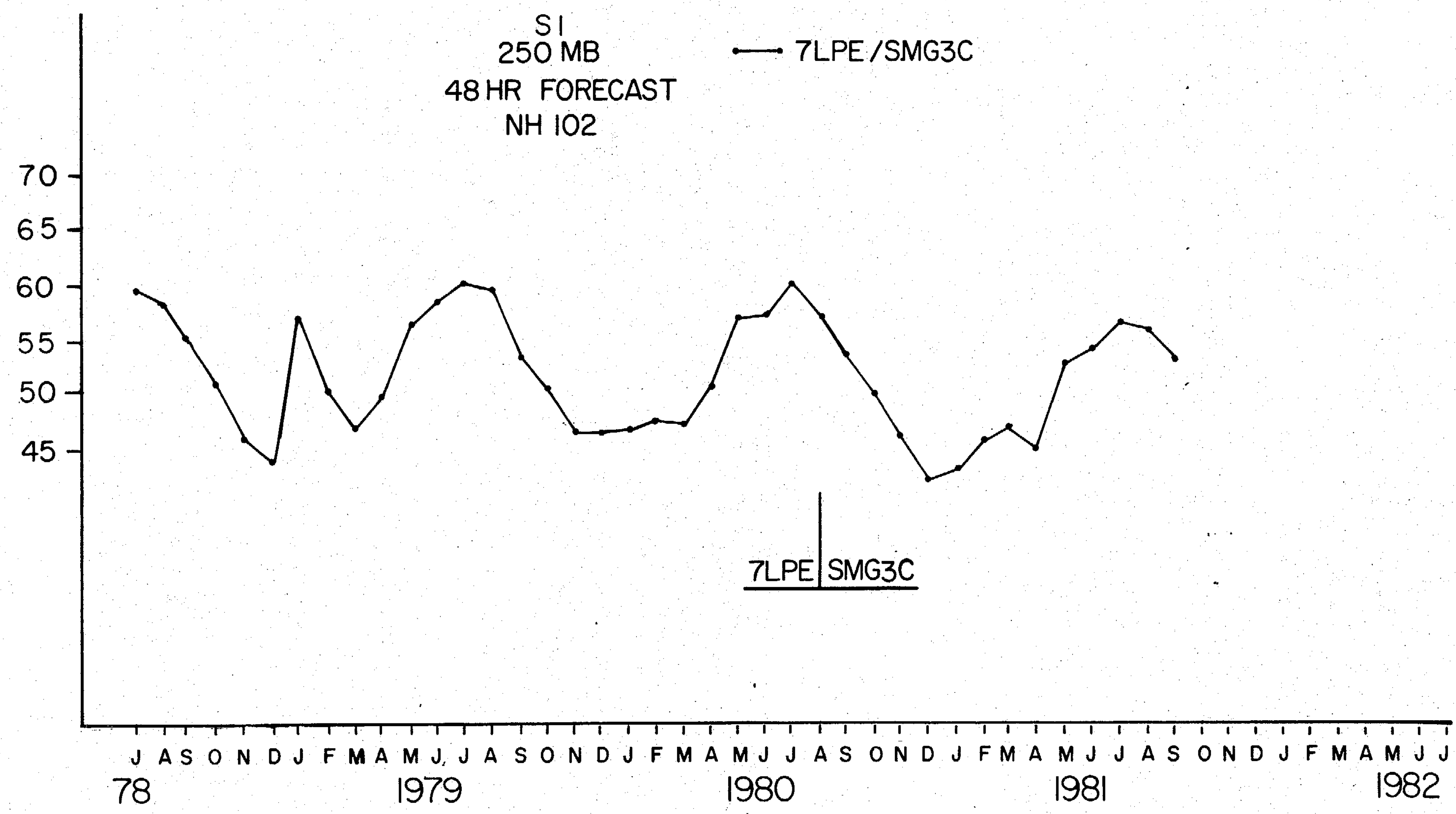

S I
250 MB
48 HR FORECAST
NH 102
7LPE/SMG3C
70
65
60
55
50
45
7LPE
SMG3C
J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J
78
1979
1980
1981
1982

FIGURE 5

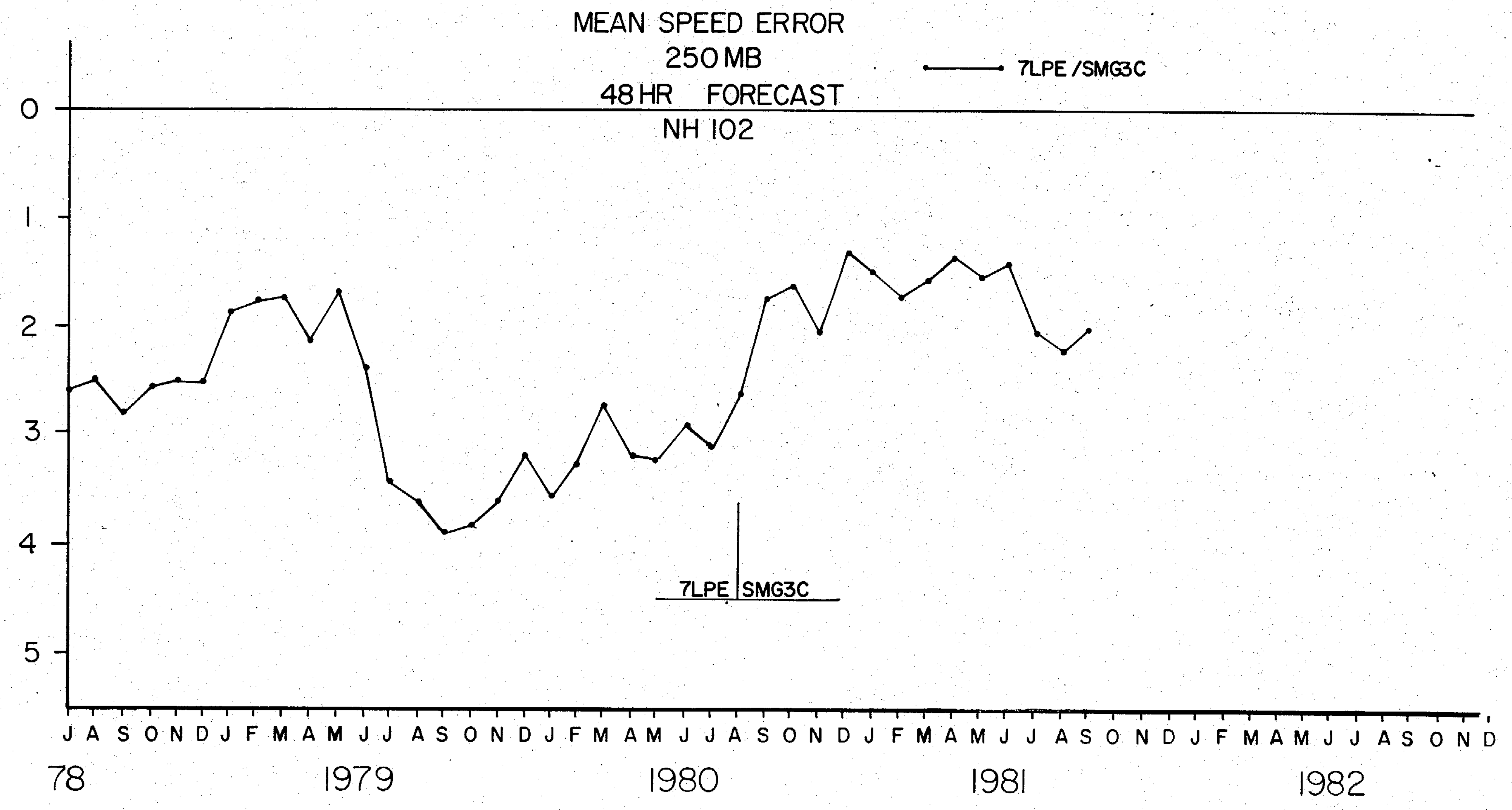
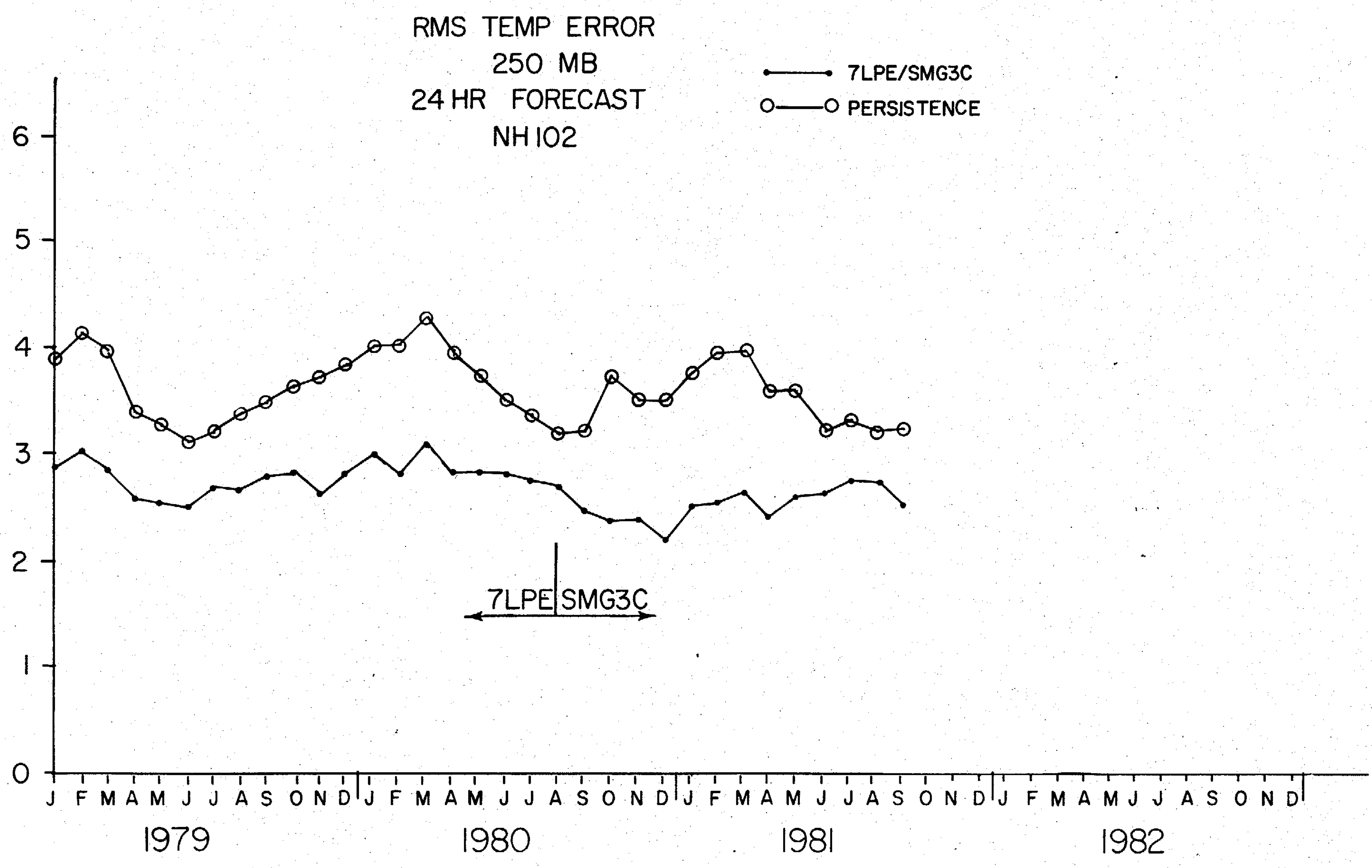

MEAN SPEED ERROR
250 MB
48 HR FORECAST
NH 102
7LPE /SMG3C
0
1
2
3
4
5
7LPE
SMG3C
J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D
78
1979
1980
1981
1982

FIGURE 6



RMS TEMP ERROR
250 MB
24 HR FORECAST
NH 102
7LPE/SMG3C
PERSISTENCE
6
5
4
3
2
1
0
7LPE
SMG3C
J F M A M J J A S O N D
J F M A M J J A S O N D
J F M A M J J A S O N D
J F M A M J J A S O N D
1979
1980
1981
1982


FIGURE 7

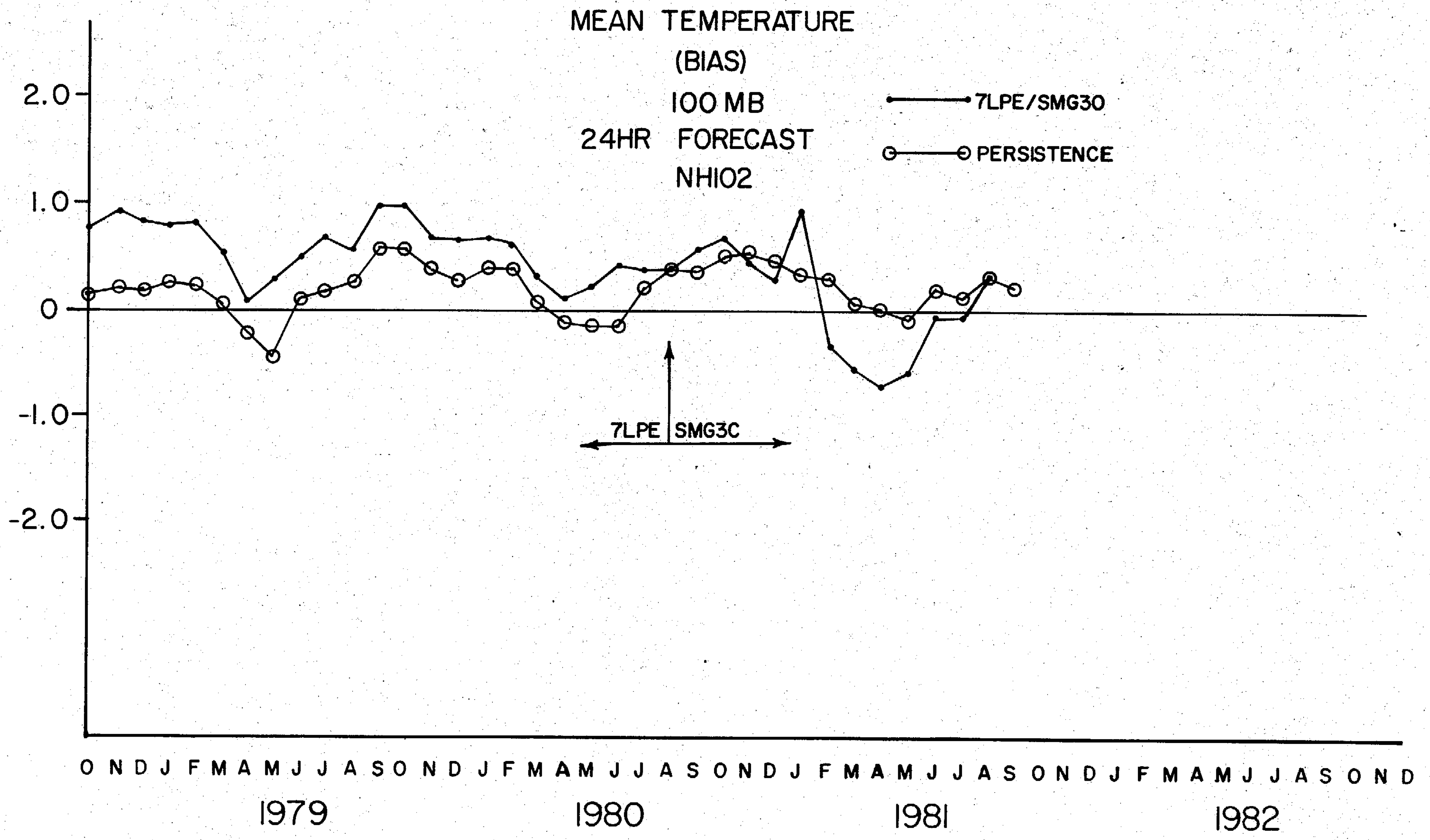

MEAN TEMPERATURE
(BIAS)
100 MB
24HR FORECAST
NH102
7LPE/SMG30
PERSISTENCE
2.0
1.0
0
-1.0
-2.0
7LPE
SMG3C
O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D
1979
1980
1981
1982

FIGURE 8

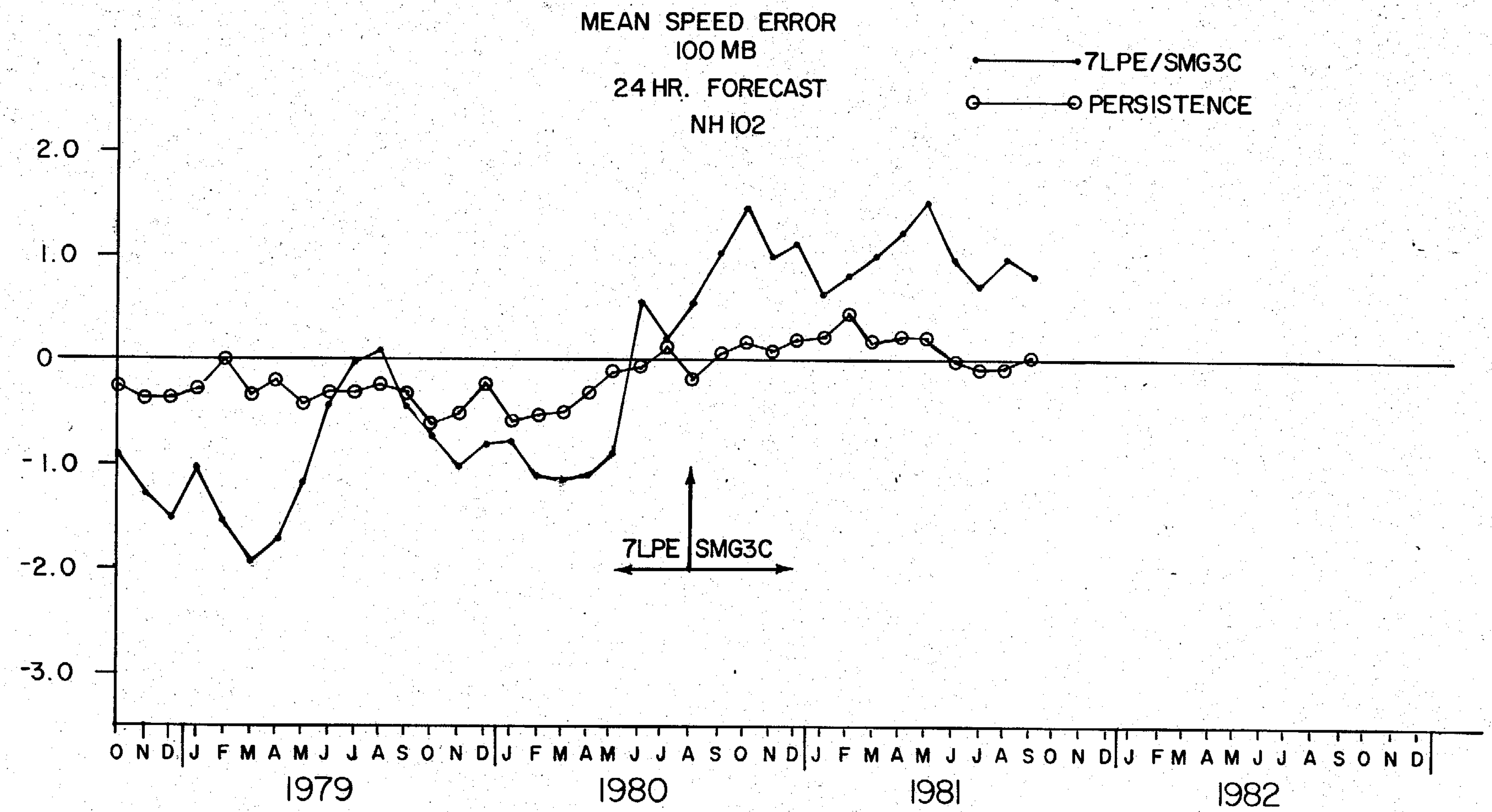
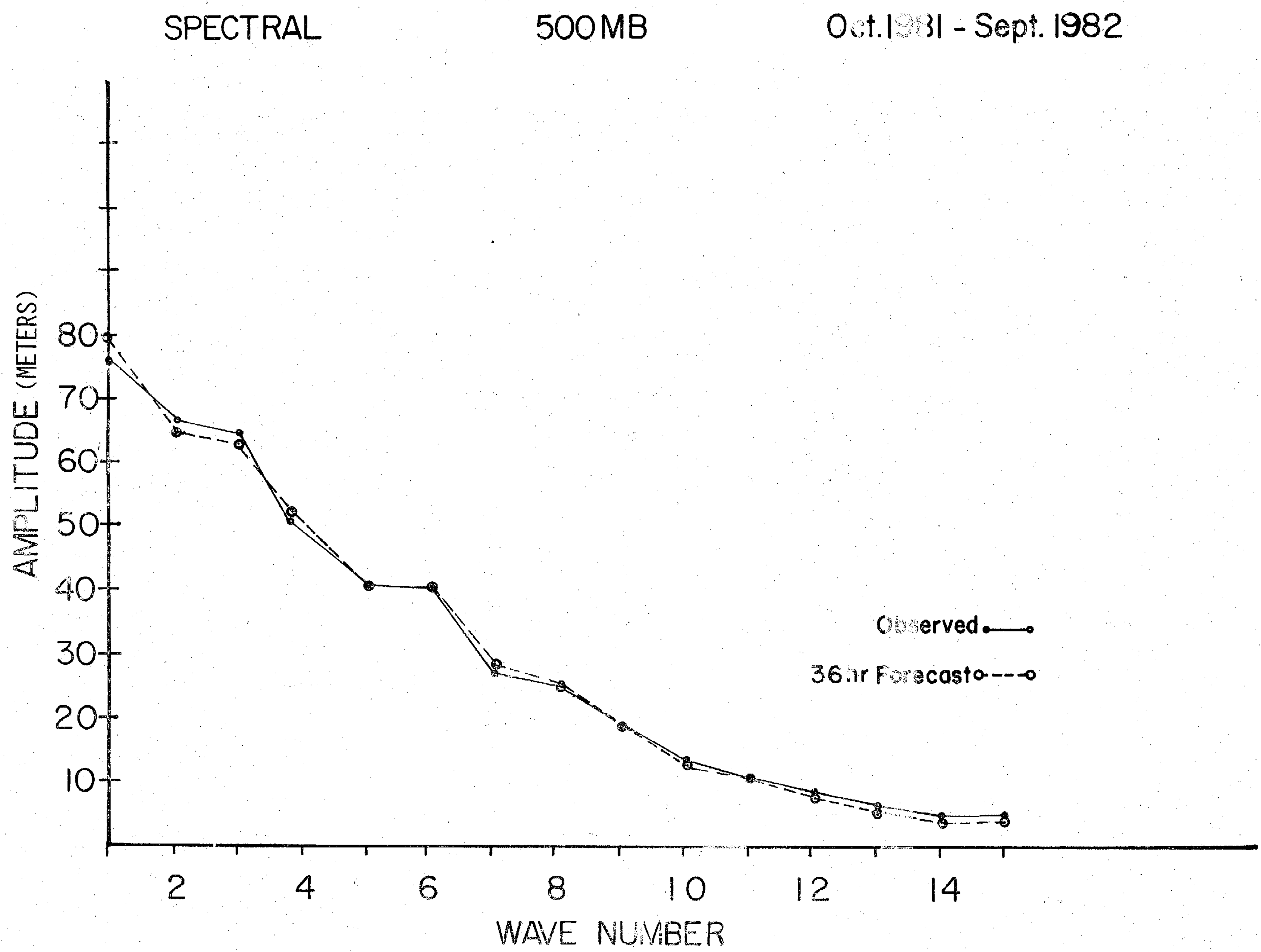
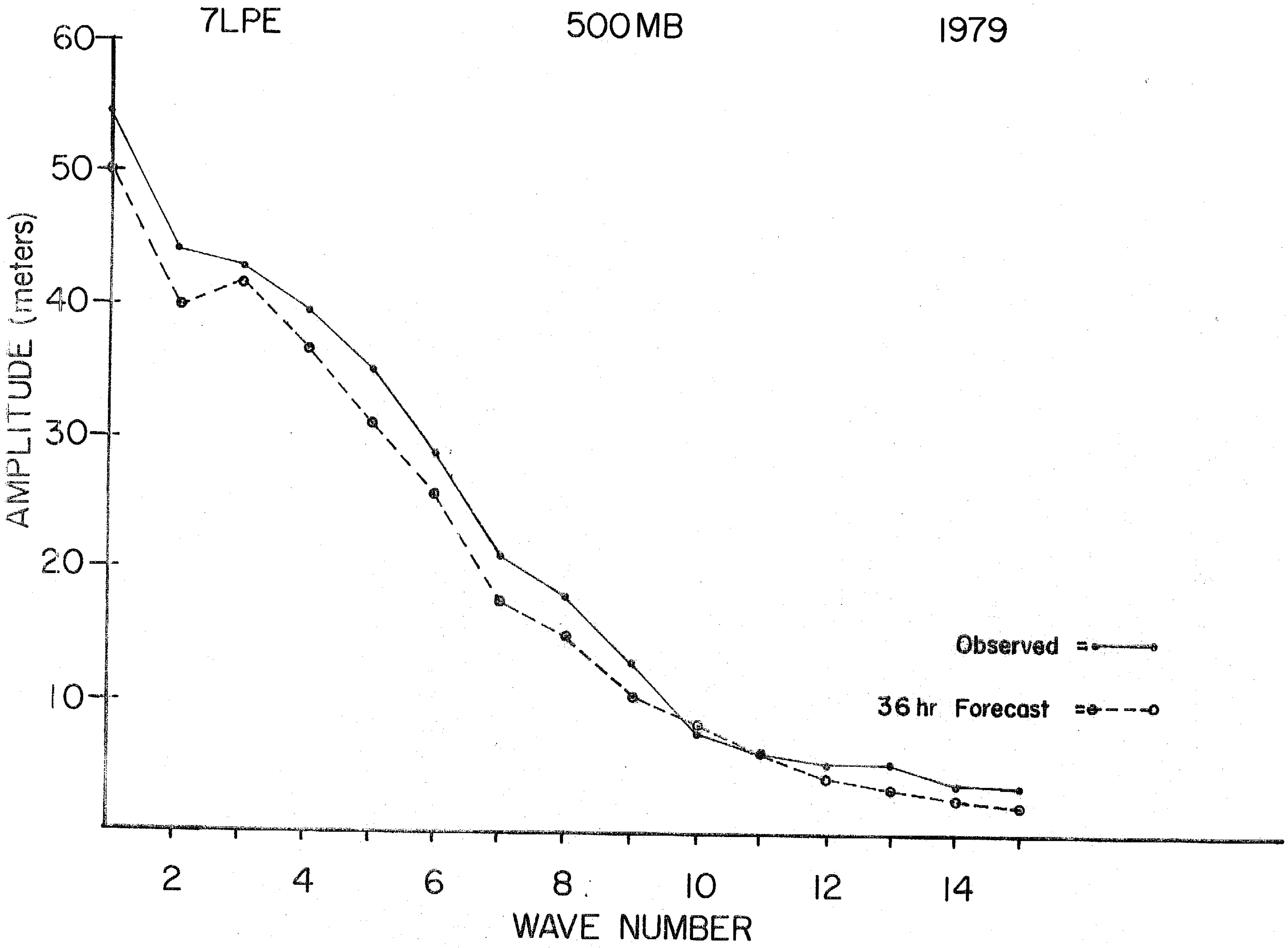

MEAN SPEED ERROR
100 MB
24 HR. FORECAST
NH 102
7LPE/SMG3C
PERSISTENCE
2.0
1.0
0
-1.0
-2.0
-3.0
7LPE
SMG3C
O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D
1979
1980
1981
1982

FIGURE 9



SPECTRAL
500MB
Oct.1981 - Sept. 1982
AMPLITUDE (METERS)
80
70
60
50
40
30
20
10
2
4
6
8
10
12
14
WAVE NUMBER
Observed
36hr Forecast


FIGURE 10


7LPE
500MB
1979
60
50
40
30
20
10
AMPLITUDE (meters)
2
4
6
8
10
12
14
WAVE NUMBER
Observed
36 hr Forecast
FIGURE 11